

基于优化全卷积神经网络的手语语义识别

王民¹, 郝静*, 要趁红, 史其琦

西安建筑科技大学信息与控制工程学院, 陕西 西安 710055

摘要 手语特征提取的传统算法仅仅依靠底层特征完成识别, 难以获得高层语义特征, 进而对手语理解产生分歧。针对这一问题, 将图像语义分析的思维引入手语识别研究中, 提出一种优化全卷积神经网络算法。采用全卷积神经网络提取手语图像的语义特征, 并通过判别随机场进行语义标注做后期平滑处理, 恢复像素间的细节信息, 从而完成手语识别。实验结果表明, 所提出的算法具有较强的稳健性, 能有效学习到语义特征。与传统算法对比分析表明, 此方法能精准的识别到手语, 其平均识别率达 97.41%。

关键词 图像处理; 图像语义; 手语识别; 全卷积神经网络; 判别随机场

中图分类号 TP391.41

文献标识码 A

doi: 10.3788/LOP55.111010

Sign Language Semantic Recognition Based on Optimized Fully Convolutional Neural Network

Wang Min¹, Hao Jing*, Yao Chenhong, Shi Qiqi

School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, Shaanxi 710055, China

Abstract The traditional algorithms for the extraction of sign language features only rely on the low-level features to realize recognition, which makes it difficult to obtain the high-level semantic features and the misunderstanding of sign language is further induced. To solve this problem, the idea of image semantic analysis is introduced into the study of sign language recognition and then an optimized fully convolutional neural network algorithm is proposed. The fully convolutional neural network is used to extract the semantic features of sign language images and the discriminative random fields for semantic annotation is used for the post-smoothing to recover the detailed information among pixels and thus the sign language recognition is completed. The experimental results show that the proposed algorithm has strong robustness and can be used to obtain the semantic features effectively. Compared with the traditional algorithms, this method can be used to identify sign language accurately with an average recognition rate of 97.41%.

Key words image processing; image semantics; sign language recognition; fully convolutional neural network; discriminative random field

OCIS codes 100.3008; 200.4260; 060.4256; 100.2960

1 引言

手语是一种借助手势来表达信息的肢体语言, 其经过手的手型、位置、朝向和运动轨迹便可表现成为一个手语词^[1]。手语识别不仅是听力障碍者的主要交流手段, 而且有效的手语识别技术将减轻听力障碍者因交流不便带来的困扰。因此, 手语识别技术具

有重要的研究意义, 并已广泛应用于人机交互, 计算机游戏, 医疗保健等领域^[2]。近年来, 研究者们通过反复的研究、实验, 提出了诸多识别方法。2014年, Sun等^[3]用 Kinect 传感器获取彩色深度图像的有效特征并通过支持向量机实现手语分类。2015年, 李萌^[4]提出用 LBP 变换与主成分分析(PCA)来对手语图像进行特征提取, 构建支持向量机分离器对手势进

收稿日期: 2018-05-03; 修回日期: 2018-06-01; 录用日期: 2018-06-08

基金项目: 国家自然科学基金(61373112)、住房和城乡建设部科学技术项目计划(2016-R2-045)、陕西省自然科学基金研究资金(2014JM8348)

* E-mail: haoanjing12@163.com

行语义识别。2016年, Huong等^[5]提出PCA对同一背景下的静态手语进行识别。2017年, 刘小建等^[6]提出使用深度信息手势图像进行多特征提取并用支持向量机作为分类器来识别手势。

有效地提取手势特征是识别的关键^[7], 这些传统的特征提取方法主要基于底层特征, 学习能力不强, 在样本量不断变大的情况下, 并不能有效提高识别率, 很难精确辨别每个像素, 进而导致人们对图像认知产生“语义鸿沟”。2017年, 王龙等^[8]借助肤色高斯模型分割出手势区域并建立卷积神经网络(CNN)模型来识别手势。这种方法有效避免了传统特征提取的局限性, 加强了算法稳定性, 但卷积神经网络算法对图像像素逐个进行语义识别时存在缺陷。2017年, Sang等^[9]设计出基于区域的全卷积神经网络(R-FCN)算法来识别复杂背景下的手势, 此方法能提取到较抽象的纹理特征, 但易丢失较小目标的细节。本文将借助图像语义分析的方法来完成精准手语识别。通过全卷积神经网络(FCN)提取手语图像的语义特征, 利用多尺度细化策略的上采样操作融合多层次的语义信息, 最终采用判别随机场(DRF)进行语义标注来优化最终结果, 实现手语语义识别。

2 算法分析

2.1 全卷积神经网络

2017年, 熊志勇等^[10]提出了全卷积神经网络(FCN)进行像素级的分类, 解决了语义级别的图像分割问题。FCN是深度学习应用在图像分割的代表作, 鉴于其在语义分析中取得的良好效果, 本文选用其来执行手语识别过程。

FCN将传统CNN中的全连接层转变为多个卷积层, CNN在卷积层之后用一个全连接层将固定长度的特征向量进行分类, 而FCN可以接受任意尺寸的输入图像, 通过反卷积层对最后一个卷积层得到的特征图进行上采样操作, 使其恢复到与输入图像相同的尺寸, 从而对每个像素都可以产生预测, 并能保留原始输入图像中的空间有效信息^[11-13]。与全连接层不同的是卷积层中的神经元只与输入数据中的一个局部区域连接^[14], 并且卷积列中的神经元共享参数。多次卷积后, 图像的分辨率越来越低, FCN经过上采样操作恢复图像的分辨率, 可输出与输入大小相同的图片, 得到每个像素的类别, 进而对图像的每个像素进行语义识别。

FCN主要包括卷积层、下采样层、上采样层和

裁剪层。卷积层是FCN的核心组成部分, 其操作原理是卷积核在图像矩阵上移动^[15-17], 卷积操作的核心是减少不必要的权值连接, 主要功能是提取图像特征信息, 并且可以平滑噪声对图像的影响。卷积层的数学表达式为

$$x_j^i = f\left(\sum_{i \in M_j} x_i^{i-1} \times k_{ij} + b_j^i\right), \quad (1)$$

式中 f 是激活函数, 在本文中采用ReLU函数作为激活函数, 其稀疏特性可以去除数据中的冗余, 最大可能保留数据的特征, 使运算速度提高, 加强了算法稳健性, b 为偏置项。

下采样过程在卷积操作之后, 二次提取特征, 这样降低了上层提取到的特征维度, 减少计算量, 保证平移不变性。降低下层输入维度, 可以有效控制因特征过多导致的过拟合风险。特征图经过降维后更能表现出图像的一般性, 更有利于分类。下采样的一般形式为

$$x_j^i = \text{down}(x_j^{i-1}), \quad (2)$$

式中 $\text{down}(\cdot)$ 是下采样函数, 其操作与卷积类似, 在经过大小为 $n \times n$ 的下采样窗口后, 图像的大小会变为输入图像的 $1/n$ 。

上采样操作可以看成是反卷积操作, 不限制输入大小, 经滑窗卷积后输出结果。其作用是将经过卷积层、下采样层后尺寸变小的图像恢复到原尺寸^[18-19], 本文采用双线性插值的方法对特征图进行上采样。

网络训练结束时, 所有层级的滤波器、上采样的参数都已固定, 此时输入不同尺寸的图片, 仅靠固定的上采样操作不能完成输出与输入尺寸一致的要求, 因此需要加入裁剪层实现输出与输入尺寸相同。

FCN数学模型为

$$\left\{ \begin{array}{l} \mathbf{X}_1 = C(x, \theta_1) \\ \mathbf{X}_2 = \text{Upsampling}(\mathbf{X}_1, \theta_2) \\ \mathbf{X}_3 = \text{Crop}[\mathbf{X}_2, \text{size}(x)] \\ y(u, v) = \text{argmax} \mathbf{X}_3(u, v, s) \\ \min J(\theta) = \frac{1}{N} \sum_{n=1}^N L(y^{(n)}, \hat{y}^{(n)}) + \lambda \cdot R(\theta) \end{array} \right., \quad (3)$$

式中: x 为输入图像; y 为输出图像; C 表示卷积操作; \mathbf{X}_1 代表最后一层卷积后得到的特征向量; Upsampling表示上采样操作; \mathbf{X}_2 代表经过上采样层将 \mathbf{X}_1 的尺寸变为与输入图像一样大小; Crop表示裁剪操作; \mathbf{X}_3 表示对 \mathbf{X}_2 裁剪; argmax表示寻求最大值的参量; $y(u, v)$ 为最后的输出; $\min J(\theta)$ 为目标

函数,参数 θ 包括 θ_1 和 θ_2 两部分, θ_1 表示各个卷积流中和分类器 softmax 函数中的参数, θ_2 为上采样中的参数;损失项 L 为 softmax 交叉熵损失函数; $R(\theta)$ 为正则项,防止过拟合。将利用随机梯度下降的方法求得各参数。

2.2 判别随机场

全卷积神经网络训练了一个端到端的网络模型,具有很强的学习能力,能够得到较准确的结果精度问题,而且不限制输入图片的尺寸,不要求图片集中所有的图片尺寸相同。但是上采样的结果对细节不敏感,会忽略像素与像素之间的关系,以及空间的一致性的问题,会影响到手势识别的精度。针对以上问题将采用 DRF 做后期处理。在完成全卷积神经网络对图像的语义特征提取后,用 DRF 进行语义标注识别,弥补以上缺点,得到最终的识别结果。

DRF 是条件随机场(CRF)对图像处理的有效扩展形式,是一种用于二维图像语义标注的概率图模型。图像建模时,DRF 模型的峰值点集代表着图像位置的集合,而边缘集合代表着相邻位置之间的联系^[20]。根据 DRF 的定义,其能量函数包括一元和二元代价函数,即

$$E = \sum \phi(c_i, x) + \sum \phi(c_i, c_j, x), \quad (4)$$

式中 c_i 为节点 i 的类标签, x 表示给定的数据。

一元代价函数 ϕ 主要是决策已知位置 i 属于某一类的趋势,不考虑邻域之间的相互影响。 ϕ 定义为

$$\phi(c_i, x) = \lg \sigma [c_i \mathbf{w}^T \mathbf{h}_i(x)], \quad (5)$$

$$\sigma(\alpha) = 1/[1 + \exp(-\alpha)], \quad (6)$$

式中 \mathbf{w} 是模型的参数, $\mathbf{h}_i(x)$ 是在每个位置 i 上对

观察数据 x 进行特征提取然后经过相应变换得到的特征向量。FCN 最终的输出为二维图像而不是概率值,因此将 FCN 的输出作为 DRF 的一元代价函数。

二元代价函数 ϕ 代表不同邻域之间的相互影响^[21],某一位置 i 上的标记通常会与其相邻位置上的标记有着很大的联系。 ϕ 定义为

$$\phi(c_i, c_j, x) = c_i c_j \mathbf{v}^T \boldsymbol{\mu}_{ij}(x), \quad (7)$$

式中 \mathbf{v} 是向量, $\boldsymbol{\mu}_{ij}$ 表示在邻域之间定义的一种新特征向量。

二元代价函数主要用于处理相邻像素的关系,采用 sobel 算子提取图像的边缘强度特征作为 DRF 的二元代价函数。Sobel 算子是一个离散的一阶差分算子,通过数字化方式来实现,用卷积核操作来提取边缘特征,可以平滑噪声并提取比较精确的边缘方向信息。

3 基于 FCN 的手势识别过程

为验证所提出的方法的可行性,自行采集 5 种包含不同语义的手势(S、T、U、V、W)作为实验的手势图像库,并与文献[4, 6, 8]中提出的方法对比分析。每种手势在不同角度,不同背景,不同亮度下共拍取 600 张,总共得到 3000 张样本。其中 2500 张作为训练样本,500 张作为测试样本。自行采集的手势如图 1 所示。为了加快训练网络的收敛性,同时保证获取足够的图像细节信息^[22],将实验中采集到的手势图像预处理为手势灰度图像,并归一化数据,得到大小为 64 pixel×64 pixel 的图像作为网络输入。

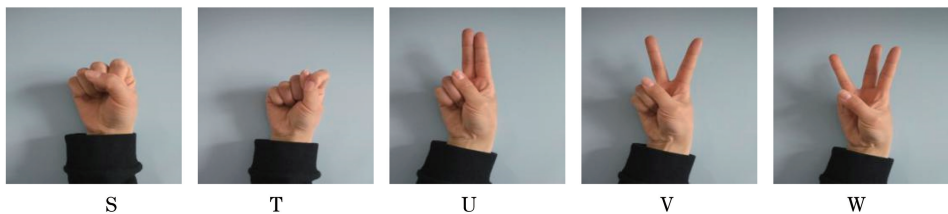


图 1 自行采集手势图像

Fig. 1 Self-acquired gesture images

全卷积神经网络结构一般由输入层、卷积层、下采样层、上采样层、裁剪层和输出层构成。本文以尺寸为 64 pixel×64 pixel 的图像作为网络输入,不再需要裁剪层,经过训练设定固定的上采样层便可以将输出图像上采样至与输入图像相同的尺寸。因为较浅的卷积层的感受域比较小,学习感知细节部分的能力较强;随着网络层数的增加,提取到的特征

越来越抽象,语义特征越来越明显^[23],感受域相对较大,适合学习较为整体的,相对更为宏观的特征。在较深的卷积层上进行反卷积操作则会丢失很多细节特征,影响手势识别的精度,因此采用多尺度细化的上采样策略,经多次上采样操作将不同尺度的语义信息融合,使网络能充分学习到有效的特征信息;采用 DRF 语义标注优化,弥补缺失较小

目标,忽略像素间的关系的缺点,有效提高识别精度。经过对不同深度、不同结构的全卷积神经网络进行测试,最终确定采用的手势识别网络结构如

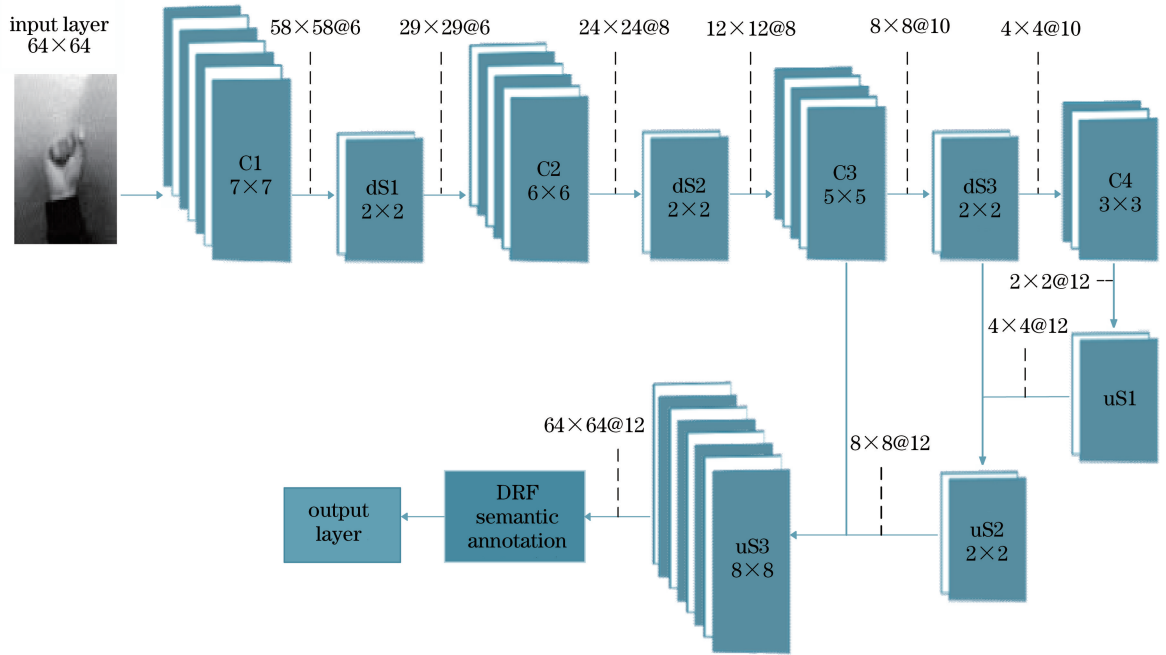


图 2 手势识别算法网络框图

Fig. 2 Network diagram of gesture recognition algorithm

表 1 全卷积神经网络参数

Table 1 Parameters of fully convolutional neural network

Parameter layer	Kernal size / (pixel×pixel)	Input / (pixel×pixel)
C1	7×7	58×58@6
dS1	2×2	29×29@6
C2	6×6	24×24@8
dS2	2×2	12×12@8
C3	5×5	8×8@10
dS3	2×2	4×4@10
C4	3×3	2×2@12
uS1	2×2	4×4@12
uS2	2×2	8×8@12
uS3	8×8	64×64@12

不同的卷积核得到不同的特征图,一个特征图代表提取到的一种特征,考虑到手势具有尺度、旋转、平移等复杂的变化,因此将依次增加每个卷积层中特征图的数量,以提高网络对图像信息的有效利用率。

网络的学习速率设置为 0.1,迭代次数设置为 200,批次大小设置为 20,通过随机梯度下降算法来训练网络,更新其参数,具体算法步骤为:

1) 输入已预处理为 64 pixel×64 pixel 的手势图像,卷积核为的 7 pixel×7 pixel 的卷积层 C1 通过卷积操作对输入图像进行特征提取,经过该层后

图 2 所示。其中 C 为卷积层,dS 为下采样层,uS 为上采样层,全卷积神经网络各层参数如表 1 所示,表中第三列@后面的数值代表特征图数量。

抽取到大小为 58 pixel×58 pixel 的 6 张特征图。

2) dS1 层对 C1 层提取到的特征图进行抽样,降低特征图的维度,刻画其平移不变性,下采样层不改变特征图的数量。经过大小为 2 pixel×2 pixel 的抽样窗口后特征图的大小变为 29 pixel×29 pixel。

3) 同步骤 1)、2) 经过 3 次卷积、2 次下采样操作后得到 12 张 2 pixel×2 pixel 的特征图。

4) 通过多尺度细化策略,将 C4 层的特征图上采样 2 倍与 dS3 层的预测信息融合进行第二次上采样 4 倍,最后与 C3 层的信息融合上采样 8 倍得到与输入图像相同大小的特征图。避免了一次性上采样至 32 倍损失很多细节信息的缺点。

5) DRF 语义标注:将 FCN 的结果作为 DRF 的一元代价函数,sobel 算子提取到的图像边缘特征作为 DRF 的二元代价函数,做平滑处理,优化 FCN 的结果,输出最终识别结果。

4 实验结果分析

对于全卷积神经网络结构,迭代次数的选择会对最终的学习效果有很大的影响。迭代次数过少会使网络无法提取到图像的有效信息而降低识别率。而迭代次数过多将会影响算法的速率。因此应用手

势数据集分析不同迭代次数与手势识别错误率的关系,如图3所示。

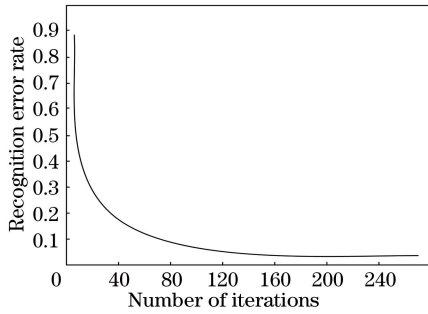


图3 手势识别错误率曲线

Fig. 3 Error rate curve of gesture recognition

从图3中可以看出,迭代次数较少时,手势识别错误率较高,识别率较低,直到迭代次数为200时,识别错误率开始趋于稳定,所以本文算法的迭代次

数设置为200,每类手势的识别率如表2所示,平均识别率为97.41%。

表2 FCN算法手势识别率

Table 2 Gesture recognition rate by FCN algorithm %

Item	Gesture category				
	S	T	U	V	W
Recognition rate	96.92	96.85	97.86	97.93	97.50

由表2可以看出,本文算法可以有效地识别到手势的语义,具有有较高的识别率,因为每类手势图像所具有的特征信息不同,所以不同手势的识别率也不同。

为验证本文算法的稳健性,给图像加上均值为0,方差分别为0.01、0.02和0.03的高斯噪声,如图4所示,在已训练好的网络上对其进行测试,识别结果如表3所示。

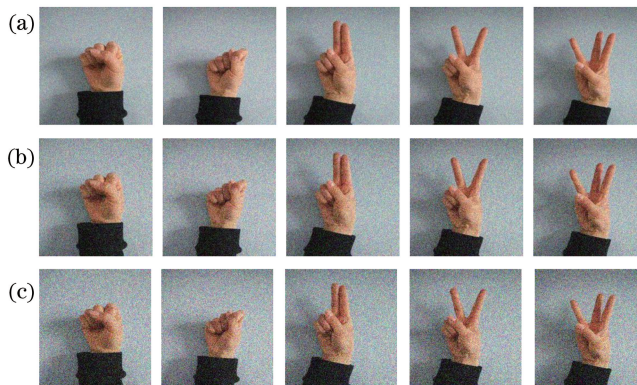


图4 加噪后的图像。(a)方差为0.01的高斯噪声;(b)方差为0.02的高斯噪声;(c)方差为0.03的高斯噪声

Fig. 4 Noised images. (a) Gaussian noise with variance of 0.01; (b) Gaussian noise with variance of 0.02; (c) Gaussian noise with variance of 0.03

表3 不同高斯噪声下的FCN算法识别率

Table 3 Recognition rate by FCN algorithm under different Gaussian noises %

Gaussian noise variance	Gesture category				
	S	T	U	V	W
0.01	96.74	96.72	97.34	97.37	97.18
0.02	96.31	96.27	97.04	97.31	96.87
0.03	96.15	96.02	96.83	97.03	96.53

由表3结果可知,加噪后的各类图像的识别率均有所下降,且随着噪声方差的增加,图像的识别率逐渐降低,但在一定程度内,图像识别率任然维持在一个较高水平。这表明该算法具有较强的稳健性,能有效抵抗噪声的干扰。

将本文算法与文献[4,6,8]提出方法的识别性能进行比较,以识别率作为评价标准,结果如表4所示。

表4 FCN与其他算法识别性能比较

Table 4 Performance comparison between FCN and other algorithms %

Algorithm	Gesture category				
	S	T	U	V	W
Algorithm in Ref. [4]	95.86	95.76	95.83	96.03	96.12
Algorithm in Ref. [6]	95.94	96.18	96.22	96.38	96.35
Algorithm in Ref. [8]	96.21	96.43	96.32	96.86	97.15
FCN algorithm	96.92	96.85	97.86	97.93	97.50

表4结果表明,在对比实验中,本文算法识别率最好,其平均识别率达到97.41%。传统的特征提取方法主要基于底层特征,不能提取到图像的抽象特征,所以识别精度不及本文算法。相对于传统的特征提取算法,卷积神经网络可以多层次的自动学习到抽象特征,通过卷积运算可以保留原始信号的空间关系,增强原始信号特征。但卷积神经网络算法的识别是图像级的识别,很难对图像的每个像素进

行语义识别。FCN 引入多尺度细化策略的上采样操作,不仅保留了细节特征,对每一个像素产生预测,而且保留了原始输入图像的空间信息,利用梯度下降学习算法的自组织能力,可以强制网络提取图像在不同特征空间下的特征并将其组合起来,使网络进行充分的特征学习。

5 结 论

构建了全卷积神经网络系统,对手语进行像素级的语义提取识别。不同于传统算法的人工手动设置,该算法通过自行学习共享权值大大减少了运行复杂度,效率大幅提升。对 5 种具有不同尺度和形变量的手势图像进行识别研究。实验表明其独特的卷积、下采样和上采样结构能准确提取到手语图像中的语义信息,使计算机对手语意义的理解更贴近人的思维,减小“语义鸿沟”,通过多尺度细化策略的上采样操作,提取更多的图像细节信息。在使用全卷积神经网络来提取图像特征后,用判别随机场进行语义标注作为后期处理方法平滑同类区域,不仅充分考虑了像素间的关系,还保留图像中的边缘信息,增加了特征识别正确率。

参 考 文 献

- [1] Jiang Y J. Research on isolated sign language recognition system based on Kinect [D]. Hefei: University of Science and Technology of China, 2015: 1-8.
江勇军. 基于 Kinect 的孤立词手语识别系统研究 [D]. 合肥: 中国科学技术大学, 2015: 1-8.
- [2] Zhang S J. Unmarked gesture recognition based on vision[M]. Jilin: Jilin University Press, 2016: 1-25.
张生军. 基于视觉的无标记手势识别 [M]. 吉林: 吉林大学出版社, 2016: 1-25.
- [3] Sun C, Zhang T, Bao B K, *et al.* Latent support vector machine for sign language recognition with Kinect[C] // IEEE International Conference on Image Processing, 2014: 4190-4194.
- [4] Li M. Algorithm of gesture recognition based on image analysis[D]. Beijing: North China University of Technology, 2015: 15-28.
李萌. 基于图像分析的手语识别技术研究 [D]. 北京: 北方工业大学, 2015: 15-28.
- [5] Huong T N T, Huu T V, Xuan T L, *et al.* Static hand gesture recognition for vietnamese sign language (VSL) using principle components analysis [C] // International Conference on Communications,

Management and Telecommunications, 2016: 138-141.

- [6] Liu X J, Zhang Y. Gesture recognition based on multi-feature and SVM classification [J]. Computer Engineering and Design, 2017, 38(4): 953-958.
刘小建, 张元. 基于多特征提取和 SVM 分类的手势识别 [J]. 计算机工程与设计, 2017, 38(4): 953-958.
- [7] Bao Z Q, Lü C G. Real-time gesture recognition based on Kinect [J]. Laser & Optoelectronics Progress, 2018, 55(3): 031008.
鲍志强, 吕辰刚. 基于 Kinect 的实时手势识别 [J]. 激光与光电子学进展, 2018, 55(3): 031008.
- [8] Wang L, Liu H, Wang B, *et al.* Gesture recognition method combining skin color models and convolution neural network [J]. Computer Engineering and Applications, 2017, 53(6): 209-214.
王龙, 刘辉, 王彬, 等. 结合肤色模型和卷积神经网络的手势识别方法 [J]. 计算机工程与应用, 2017, 53(6): 209-214.
- [9] Sang N, Ni Z H. Gesture recognition based on R-FCN in complex scenes [J]. Journal of Huazhong University of Science and Technology(Nature Science Edition), 2017, 45(10): 54-58.
桑农, 倪子涵. 复杂场景下基于 R-FCN 的手势识别 [J]. 华中科技大学学报(自然科学版), 2017, 45(10): 54-58.
- [10] Xiong Z Y, Zhang G F, Wang J Q. Multi-scale feature extract for image sematic segmentation [J]. Journal of South-Central University for Nationalities (Natural Science Edition), 2017, 36(3): 118-124.
熊志勇, 张国丰, 王江晴. 基于多尺度特征提取的图像语义分割 [J]. 中南民族大学学报(自然科学版), 2017, 36(3): 118-124.
- [11] Jiao L C, Zhao J, Yang S Y, *et al.* Deep learning, optimization and identification[M]. Beijing: Tsinghua University Press, 2017: 77-80.
焦李成, 赵进, 杨淑媛, 等. 深度学习, 优化与识别 [M]. 北京: 清华大学出版社, 2017: 77-80.
- [12] Maggiori E, Tarabalka Y, Charpiat G, *et al.* Fully convolutional neural networks for remote sensing image classification [C] // Geoscience and Remote Sensing Symposium, 2016: 5071-5074.
- [13] Fang X, Wang G H, Yang H C, *et al.* High resolution remote sensing image classification combining with mean-shift segmentation and fully convolution neural network [J]. Laser & Optoelectronics Progress, 2018, 55(2): 022802.

- 方旭, 王光辉, 杨化超, 等. 结合均值漂移分割与全卷积神经网络的高分辨遥感影像分类[J]. 激光与光电子学进展, 2018, 55(2): 022802.
- [14] Milletari F, Navab N, Ahmadi S A. V-Net: fully convolutional neural networks for volumetric medical image segmentation[C]//International Conference on 3D Vision, 2016: 565-571.
- [15] Cao X W, Bo H. Study on gesture recognition based on CNN [J]. Microcomputer & its Applications, 2016, 35(9): 55-57.
- 操小文, 薄华. 基于卷积神经网络的手势识别研究[J]. 微型机与应用, 2016, 35(9): 55-57.
- [16] Kang K, Wang X. Fully Convolutional Neural Networks for Crowd Segmentation [J]. Computer Science, 2014, 49(1): 25-30.
- [17] Li S M, Lei G Q, Fan R. Depth map super-resolution reconstruction based on convolutional neural networks [J]. Acta Optica Sinica, 2017, 37(12): 1210002.
- 李素梅, 雷国庆, 范如. 基于卷积神经网络的深度图超分辨率重建 [J]. 光学学报, 2017, 37(12): 1210002.
- [18] Yi M, Sui L C. Aerial image semantic classification method based on improved full convolution neural network [J]. Computer Engineering, 2017, 43(10): 216-221.
- 易盟, 隋立春. 基于改进全卷积神经网络的航拍图像语义分类方法[J]. 计算机工程, 2017, 43(10): 216-221.
- [19] Luo B, Gao W, Tang J, *et al.* Learning corner regression-based fully convolutional neural network for license plate localization in complex scene [J]. Journal of Data Acquisition & Processing, 2016, 31(1): 65-72.
- 罗斌, 郜伟, 汤进, 等. 复杂环境下基于角点回归的全卷积神经网络的车牌定位[J]. 数据采集与处理, 2016, 31(1): 65-72.
- [20] Pan C, Wang X J. Summary of algorithms for detecting man-made information in natural background [J]. Optical Technique, 2008, 34(6): 881-885.
- 潘聪, 王向军. 自然背景中人造信息的检测算法综述 [J]. 光学技术, 2008, 34(6): 881-885.
- [21] Ding H, Zhang X F. Connected handwritten and printed text discrimination in uneven lighted images [J]. Computer Engineering and Design, 2012, 33(12): 4634-4638.
- 丁红, 张晓峰. 非均匀光照图像中粘连手写体和印刷体的辨别 [J]. 计算机工程与设计, 2012, 33(12): 4634-4638.
- [22] Zeng W J, Qin F, Si Y. Sign language recognition system based on artificial neural network [J]. China Computer & Communication, 2017(11): 143-144.
- 曾维佳, 秦放, 思雨. 基于人工神经网络的手语识别系统 [J]. 信息与电脑, 2017(11): 143-144.
- [23] Xin P, Xu Y L, Tang H, *et al.* Fast airplane detection based on multi-layer feature fusion of fully convolutional networks [J]. Acta Optica Sinica, 2018, 38(3): 031500.
- 辛鹏, 许悦雷, 唐红, 等. 全卷积网络多层特征融合的飞机快速检测 [J]. 光学学报, 2018, 38(3): 031500.